WHAT IS CLAIMED IS:

1. A method for isolating a collection of polynucleotide sequences corresponding to accessible regions of cellular chromatin, the method comprising:

(a) treating cellular chromatin with a probe, wherein the probe reacts with accessible regions of cellular chromatin;

(b) fragmenting the treated chromatin to produce unmarked polynucleotides and marked polynucleotides, wherein each marked polynucleotide comprises one or more sites that have reacted with the probe; and

(c) isolating the marked polynucleotides.

2. The method of claim 1, wherein, subsequent to step (a), the treated chromatin is deproteinized.

3. The method of claim 1, wherein the method comprises cloning each marked polynucleotide.

4. The method of claim 2, wherein fragmentation is by restriction enzyme digestion.

5. The method of claim 4, wherein the probe is a DNA methylase and each marked polynucleotide is methylated.

6. The method of claim 5, wherein the DNA methylase is *E. coli* dam methylase.

7. The method of claim 5, wherein the restriction enzyme is a methylation-sensitive restriction enzyme.

8. The method of claim 5, wherein each methylated polynucleotide is isolated according to size.

9. The method of claim 5, wherein each methylated polynucleotide is isolated by binding to an antibody specific for methylated DNA.

10. The method of claim 1, wherein the treating is performed on isolated chromatin.

11. The method of claim 1, wherein the treating is performed on isolated nuclei

12. The method of claim 1, further comprising sequencing one or more of the marked polynucleotides.

13. The method of claim 1, wherein the probe is an enzyme or an antibody.

14. The method of claim 1, wherein the probe is a chemical.

15. The method of claim 1, further comprising the step of isolating the unmarked polynucleotides comprising non-accessible regions of cellular chromatin.

16. A method for isolating a collection of polynucleotides corresponding to accessible regions of cellular chromatin, the method comprising:

(a) treating cellular chromatin with a methylase to generate methylated chromatin;

(b) deproteinizing the methylated chromatin to form deproteinized chromatin;

(c) digesting the deproteinized chromatin with a methylation-dependent restriction enzyme to produce a collection of restriction fragments, wherein the collection comprises methylated polynucleotides and non-methylated polynucleotides; and

(d) isolating a collection of non-methylated polynucleotides;

whereby the termini of the non-methylated polynucleotides correspond to accessible regions of cellular chromatin.

17. The method of claim 16, wherein, in step (a), cellular chromatin is treated with *E. coli* dam methylase.

18. The method of claim 17, wherein, in step (c), the methylation-dependent restriction enzyme is DpnI.

19. The method of claim 16, wherein, in step (d), the non-methylated polynucleotides are collected according to size.

20. The method of claim 16, further comprising determining the nucleotide sequences at the termini of the non-methylated polynucleotides.

21. A method for isolating a collection of polynucleotides corresponding to accessible regions of cellular chromatin, the method comprising:

(a) treating cellular chromatin with a methylase to generate methylated chromatin;

(b) deproteinizing the methylated chromatin to form deproteinized chromatin;

(c) digesting the deproteinized chromatin with a methylation-dependent restriction enzyme to produce a collection of restriction fragments, wherein the collection comprises methylated polynucleotides and non-methylated polynucleotides; and

(d) isolating a collection of methylated polynucleotides,

whereby the methylated polynucleotides correspond to accessible regions of cellular chromatin.

22. The method of claim 21, wherein, in step (a), cellular chromatin is treated with *E. coli* dam methylase.

23. The method of claim 22, wherein, in step (c), the methylation-dependent restriction enzyme is DpnI.

24. The method of claim 21, wherein, in step (d), the methylated polynucleotides are collected according to size.

25. The method of claim 21, wherein, in step (d), the methylated polynucleotides are collected by binding to an antibody specific for methylated DNA.

26. The method of claim 21, further comprising the step of determining the nucleotide sequences of the methylated polynucleotides.

27. A method for isolating a collection of polynucleotides corresponding to accessible regions of cellular chromatin, the method comprising:

(a) treating cellular chromatin with a nuclease; and

(b) isolating a collection of polynucleotide fragments released by nuclease treatment; wherein the released polynucleotide fragments are derived from accessible regions of cellular chromatin.

28. The method of claim 27, wherein the nuclease is DNase I.

29. The method of claim 27, further comprising determining the nucleotide sequences of the collected fragments.

30. The method of claim 27, further comprising collecting the cellular chromatin remaining after nuclease treatment to generate a collection of polynucleotides comprising non-accessible regions of cellular chromatin.

31. A method for isolating a collection of polynucleotides corresponding to regulatory regions of cellular chromatin, the method comprising:

(a) treating cellular chromatin with a methylation-sensitive enzyme that cleaves at unmethylated CpG sequences; and

(b) isolating a collection of short polynucleotide fragments released by enzyme treatment, wherein the polynucleotide fragments are derived from regulatory regions of cellular chromatin.

32. The method of claim 31, wherein the methylation-sensitive enzyme is a restriction enzyme.

33. The method of claim 32, wherein the methylation-sensitive restriction enzyme is HpaII.

34. The method of claim 31, wherein, in step (a), cellular chromatin is deproteinized prior to enzyme treatment.

35. The method of claim 31, wherein collecting comprises collecting the polynucleotide fragments according to size.

36. The method of claim 31, further comprising determining the nucleotide sequences of the collected polynucleotide fragments.

37. The method of claim 31, further comprising collecting the cellular chromatin remaining after restriction enzyme treatment to generate a collection of polynucleotides comprising non-accessible regions of cellular chromatin.

38. A method for isolating a collection of polynucleotides corresponding to regulatory regions of cellular chromatin, the method comprising:

(a) selectively cleaving AT-rich sequences of cellular DNA; and

(b) isolating a collection of large polynucleotide fragments released by the treatment; wherein the large polynucleotide fragments comprise regulatory regions.

39. The method of claim 38, wherein the selective cleavage is performed by an enzyme.

40. The method of claim 39, wherein the enzyme is a restriction enzyme selected from the group consisting of Mse I, Tsp509 I, Ase I, Dra I, Pac I, Psi I, Ssp I and Swa I.

41. The method of claim 38, further comprising determining the nucleotide sequences of the collected polynucleotide fragments.

42. A method for isolating a collection of polynucleotides corresponding to regulatory regions of cellular chromatin, the method comprising:

(a) selectively cleaving AT-rich sequences in cellular DNA to form a mixture of methylated and unmethylated fragments enriched in CpG islands; and

(b) isolating the unmethylated fragments from the methylated fragments to obtain a collection of unmethylated fragments enriched in CpG islands, wherein the unmethylated fragments are derived from regulatory regions of cellular chromatin.

43. The method of claim 42, wherein separating comprises contacting the methylated fragments with an antibody specific for the methylated fragments to form an immunoprecipitate that is separated from the unmethylated fragments.

44. The method of claim 42, wherein the selective cleaving employs a restriction enzyme selected from the group consisting of Mse I, Tsp509 I, Ase I, Dra I, Pac I, Psi I, Ssp I and Swa I.

45. The method of claim 42, further comprising determining the nucleotide sequences of the unmethylated fragments.

46. The method of claim 42, further comprising collecting the methylated fragments comprising non-accessible regions.

47. A method for isolating a collection of polynucleotides corresponding to regulatory regions of cellular chromatin, the method comprising:

(a) fragmenting chromatin;

(b) contacting the fragments with an antibody that specifically binds to acetylated histones, thereby forming an immunoprecipitate enriched in polynucleotides corresponding to regulatory regions; and

(d) collecting the polynucleotides from the immunoprecipitate.

48. The method of claim 47, wherein the antibody specifically binds to acetylated histone H3.

49. The method of claim 47, further comprising cross-linking histones to the DNA in chromatin prior to fragmenting.

50. The method of claim 49, wherein crosslinking comprises exposing the sample of cellular chromatin to ultraviolet radiation.

51. The method of claim 49, wherein crosslinking comprises contacting the sample of cellular chromatin with a chemical cross-linking agent.

52. The method of claim 51, wherein the chemical crosslinking agent is selected from the group consisting of formaldehyde and psoralens.

53. The method of claim 47, further comprising determining the nucleotide sequences of the collected polynucleotides.

54. A method for mapping accessible regions of cellular chromatin relative to a gene of interest, the method comprising:

(a) reacting cellular chromatin with a probe to generate chromatin-associated DNA fragments, wherein the DNA fragments comprise, at their termini, sites of probe reaction which identify accessible regions of cellular chromatin;

(b) attaching an adapter polynucleotide to the termini generated by the probe to generate adapter-ligated fragments; and

(c) amplifying the adapter-ligated fragments in the presence of a first primer that is complementary to the adapter and a second primer that is complementary to a segment of the gene of interest to form one or more amplified products, wherein the size of an amplified product is a measure of the distance between the segment of the gene to which the second primer binds and a terminus generated by the probe;

thereby mapping accessible regions of cellular chromatin relative to the gene of interest.

55. The method of claim 54, wherein the probe is an enzymatic probe.

56. The method of claim 55, wherein the enzymatic probe is a nuclease.

57. The method of claim 54, wherein a plurality of second primers, each complementary to a segment of a different gene of interest, are used to generate a plurality of amplification products.

58. The method of claim 54, further comprising determining the nucleotide sequences of the amplified products.

59. A method for generating a library of polynucleotides whose sequences correspond to accessible regions of cellular chromatin, the method comprising:

(a) reacting cellular chromatin with a probe to generate chromatin-associated DNA fragments, wherein the DNA fragments comprise, at their termini, sites of probe reaction which identify accessible regions of cellular chromatin;

(b) attaching a first adapter polynucleotide to the termini generated by the probe to generate adapter-ligated fragments;

(c) digesting the adapter-ligated fragments with a restriction enzyme to generate a population of digested fragments, wherein the population comprises digested fragments having a first end that comprises the first adapter and a second end formed via the activity of the restriction enzyme;

(d) contacting the digested fragments with a primer complementary to the first adapter under conditions wherein the primer is extended to generate a plurality of extension products, each comprising a first end that comprises the first adapter and a second end that can be attached to a second adapter polynucleotide;

(e) joining the second adapter to the second end of each of the plurality of extension products to form a plurality of modified fragments, each of which comprises the first and second adapters at its first and second end, respectively;

(f) amplifying the plurality of modified fragments in the presence of primers complementary to the sequences of the first and second adapters to generate a population of amplified products comprising sequences corresponding to accessible regions of cellular chromatin; and

(g) inserting the population of amplified products into a selected vector;

thereby generating a library of polynucleotides whose sequences correspond to accessible regions of cellular chromatin.

60. The method of claim 59, wherein the probe is an enzymatic probe.

61. The method of claim 60, wherein the enzymatic probe is a nuclease.

62. The method of claim 61, wherein the nuclease is DNase I or micrococcal nuclease.

63. The method of claim 59, wherein the probe is a chemical probe.

64. The method of claim 59, wherein the restriction enzyme has a four-nucleotide recognition sequence.

65. The method of claim 59, further comprising determining the nucleotide sequences of the amplified products.

66. A library of polynucleotides comprising polynucleotides that correspond to accessible regions of cellular chromatin, wherein the polynucleotides are obtained according to claim 3.

67. A plurality of libraries each library comprising polynucleotides, the polynucleotides in each library corresponding to accessible regions of cellular chromatin, wherein

(a) the polynucleotides in each library are obtained according to the method of claim 3; and

(b) the polynucleotides in different libraries correspond to accessible regions in different chromatin samples.

68. The libraries of claim 67, wherein the different cellular chromatin samples are obtained from cells at different stages of development.

69. The libraries of claim 67, wherein the different cellular chromatin samples are obtained from different tissues.

70. The libraries of claim 67, wherein the different cellular chromatin samples are obtained from healthy and diseased cells.

71. The libraries of claim 67, wherein the different cellular chromatin samples are obtained from infected and uninfected cells.

72. A database comprising a collection of polynucleotide sequences corresponding to accessible regions of cellular chromatin, wherein the sequences are obtained according to the method of claim 12.

73. A method for comparing a plurality of cell populations, the method comprising the steps of:

(a) obtaining a database of polynucleotide sequences according to claim 72 for each population;

(b) comparing the sequences in two or more of the databases; and

(c) determining sequences which are unique to at least one of the cell populations.

74. The method according to claim 73, wherein the cell populations are at different developmental stages.

75. The method according to claim 73, wherein the cell populations are from different tissues.

76. The method according to claim 73, wherein at least one of the cell populations comprises diseased cells and at least one other cell population comprises counterpart normal cells.

77. The method according to claim 73, wherein at least one of the cell populations is infected with a pathogen and at least one other cell population comprises counterpart uninfected cells.

78. A method for analyzing polynucleotide sequences, comprising:

(a) providing a database comprising a plurality or polynucleotide sequences corresponding to accessible regions of cellular chromatin, wherein the polynucleotide sequences are organized in collections corresponding to accessible regions from different samples of cellular chromatin;

(b) selecting two or more of the collections for comparison;

(c) determining whether one or more of the collections being compared includes a common polynucleotide sequence;

(d) determining whether one or more of the collections being compared includes a unique polynucleotide sequence; and

(e) displaying the common or unique sequences.

79. The method of claim 78, wherein step(e) comprises displaying both common and unique sequences.

80. The method of claim 78, wherein one or more of the collections are obtained according to the method of claim 1.

81. The method of claim 78, wherein each of the plurality of collections of polynucleotide sequences are stored on a computer-readable medium and one or more of the steps are performed by a computer.

82. The method of claim 78, wherein the different samples are obtained from cells at different developmental stages.

83. The method of claim 78, wherein the different samples are obtained from different types of tissue.

84. The method of claim 78, wherein at least one of the different samples is obtained from diseased cells and at least one of the other samples is obtained from counterpart normal cells.

85. The method of claim 78, wherein at least one of the different samples is obtained from cells infected with a pathogen and at least one of the other samples is obtained from counterpart uninfected cells.

86. The method of claim 78, wherein the different samples are obtained from different cell populations that express a gene of interest at different levels.

87. The method of claim 86, wherein two or more polynucleotides sequences are common to two or more of the collections.

88. The method of claim 86, wherein at least one cell population expresses the gene of interest at higher levels relative to at least one other cell population.

89. The method of claim 87, wherein comparing comprises determining whether the collections of polynucleotides from samples obtained from one or more cell population expressing the gene of interest at a high level, compared to at least one other sample, include a unique sequence.

90. The method of claim 87, wherein

(a) the plurality of collections comprises a first collection of polynucleotide sequences corresponding to a sample obtained from a first population of cells expressing the gene of interest and a second collection of polynucleotide sequences corresponding to a sample obtained from a second population of cells expressing the gene of interest at a lower level, compared to the first population of cells; and

(b) comparing comprises determining whether the first collection includes a unique polynucleotide sequence, such a sequence potentially representing a positive regulatory sequence.

91. The method of claim 87, wherein

(a) the plurality of collections comprises a first collection of polynucleotide sequences corresponding to a sample obtained from a first population of cells expressing the gene of interest and a second collection of polynucleotide sequences corresponding to a sample obtained from a second population of cells expressing the gene of interest at a lower level, compared to the first population of cells; and

(b) comparing comprises determining whether the second collection includes a unique polynucleotide sequence, such a sequence potentially representing a negative regulatory sequence.

92. The method of claim 78, wherein the unique sequence is a polymorphic sequence.

93. The method of claim 92, wherein the polymorphism is selected from the group consisting of SNPs, haplotypes and oligonucleotide repeat expansions.

94. A method for analyzing polynucleotide sequences, comprising:

(a) obtaining a collection of polynucleotide sequences, the collection comprising a plurality of polynucleotide sequences corresponding to accessible regions of

cellular chromatin in a sample, wherein the collection is obtained according to the method of claim 1;

(b) comparing the collection of polynucleotide sequences with one or more known sequences.

95. The method of claim 94, wherein the collection is stored on a computer-readable medium and comparing is performed with a computer.

96. The method of claim 94, wherein the one or more known sequences is a known regulatory sequence.

97. The method of claim 96, wherein the regulatory sequence is selected from the group consisting of a locus control sequence, an enhancer sequence, a promoter sequence, a boundary element sequence and a matrix attachment sequence.

98. The method of claim 94, wherein

(a) the one or more known sequences is the genome sequence of the organism from which the sample was obtained; and

(b) comparing comprises determining the location of one or more of the polynucleotide sequences within the collection relative to the genome sequence.

99. The method of claim 94, wherein

(a) the one or more known sequences are sequences that include a single nucleotide polymorphism; and

(b) comparing comprises determining whether the sequences in the collection include the polymorphisms of the known sequences.

100. A computer system for analyzing polynucleotide sequences, comprising:

(a) a memory;

(b) a system bus; and

(c) a processor programmed to:

(i) compare one or more polynucleotide sequences from each of a plurality of collections of polynucleotide sequences, wherein each collection comprises a

plurality of polynucleotide sequences corresponding to accessible regions of cellular chromatin, different collections comprising polynucleotide sequences that correspond to accessible regions for different samples of cellular chromatin;

(ii) identify one or more polynucleotides unique or common to at least one of the plurality of collections; and

(iii) display the identified polynucleotide sequence(s).

101. A computer system, comprising:

(a) a database comprising sequence records that include an identifier that identifies one or more projects to which each of the sequence records belong, each of the projects comprising

(i) comparing a plurality of polynucleotide sequences from each of a plurality of collections of polynucleotide sequences, wherein each collection comprises a plurality of polynucleotide sequences corresponding to accessible regions of cellular chromatin, different collections comprising polynucleotide sequences that correspond to accessible regions for different samples of cellular chromatin, and

(ii) identifying one or more polynucleotide sequences that are unique or common to at least one of the plurality of collections; and

(b) a user interface that permits a user to selectively view information concerning the one or more projects.

102. The computer system of claim 101, wherein the different samples are obtained from cells at different developmental stages.

103. The computer system of claim 101, wherein the different samples are obtained from different types of tissues.

104. The computer system of claim 101, wherein at least one of the samples is obtained from diseased cells and at least one of the samples is obtained from counterpart normal cells.

105. The computer system of claim 101, wherein at least one of the samples is obtained from infected cells and at least one of the samples is obtained from counterpart uninfected cells.

106. The computer system of claim 101, wherein the different samples are obtained from different cell populations that express a gene of interest at different levels.

107. A computer system, comprising:

(a) a database comprising sequence records that include an identifier that identifies one or more projects to which each of the sequence records belong, each of the projects comprising

(i) comparing a plurality of polynucleotide sequences from each of a plurality of collections of polynucleotide sequences, wherein each collection comprises a plurality of polynucleotide sequences corresponding to accessible regions of cellular chromatin, different collections comprising polynucleotide sequences that correspond to accessible regions for different samples of cellular chromatin, and

(ii) identifying one or more polynucleotide sequences unique or common to at least some of the plurality of collections; and

(b) a user interface that

(i) permits a user to input identifying information specifying which of the polynucleotide sequences of the plurality of collections are to be compared; and

(ii) displays the identified polynucleotide(s).

108. A computer system for analyzing polynucleotide sequences, comprising:

(a) a memory;

(b) a system bus; and

(c) a processor operatively disposed to

(i) compare a collection of polynucleotide sequences corresponding to accessible regions of cellular chromatin in a sample with one or more known sequences to assess sequence similarity between one or more of the polynucleotide sequences within the collection and the one or more known sequences; and

(ii) display information concerning the sequence similarity between the one or more of the polynucleotide sequences within the collection and the one or more known sequences.

109. A computer system, comprising:

(a) a database comprising sequence records that include an identifier that identifies one or more projects to which each of the sequence records belong, each of the projects comprising comparing a collection of polynucleotide sequences corresponding to accessible regions of cellular chromatin in a sample with one or more known sequences to assess sequence similarity between one or more polynucleotide sequences within the collection and the one or more known sequences; and

(b) a user interface that permits a user to selectively view information concerning the one or more projects.

110. The computer system of claim 109, wherein

(a) the one or more known sequences comprise a known regulatory sequence; and

(b) the information comprises information on whether the collection of polynucleotide sequences includes the one or more known regulatory sequences.

111. The computer system of claim 109, wherein

(a) the one or more known sequences is the genome sequence of the organism from which the sample was obtained; and

(b) the user interface permits the viewer to view the location of one or more of the polynucleotide sequences of the collection relative to the genome sequence.

112. A computer system, comprising:

(a) a database comprising sequence records that include an identifier that identifies one or more projects to which each of the sequence records belong, each of the projects comprising comparing a collection of polynucleotide sequences corresponding to accessible regions of cellular chromatin in a sample with one or more known sequences to assess sequence similarity between one or more polynucleotide sequences within the collection and the one or more known sequences; and

(b) a user interface that

(i) permits a user to input identifying information specifying which of the polynucleotide sequences within the collections are to be compared; and

(ii) displays information regarding sequence similarity between the one or more polynucleotides sequences and the one or more known sequences.

113. A computer-readable medium comprising program instructions for analyzing polynucleotide sequences by performing the following:

(a) providing or receiving a plurality of collections of polynucleotide sequences, each collection comprising a plurality of polynucleotide sequences corresponding to accessible regions of cellular chromatin, different collections comprising accessible regions for different samples of cellular chromatin;

(b) identifying one or more polynucleotide sequences that are unique to at least one of the plurality of collections;

(c) identifying one or more polynucleotide sequences that are common to two or more of the plurality of collections; and

(d) displaying information concerning any identified polynucleotide sequence(s).

114. The computer-readable medium of claim 113, further comprising program instructions for comparing one or more sequences from each of the plurality of collections.

115. The computer-readable medium of claim 114, further comprising program instructions for comparing each of the sequences from each of the plurality of collections.

116. The computer-readable medium of claim 113, wherein the different samples are obtained from cells at different developmental stages.

117. The computer system of claim 113, wherein the different samples are obtained from different types of tissues.

118. The computer system of claim 113, wherein at least one of the samples is obtained from diseased cells and at least one of the samples is obtained from counterpart normal cells.

119. The computer system of claim 113, wherein at least one of the samples is obtained from infected cells and at least one of the samples is obtained from counterpart uninfected cells.

120. A computer-readable medium comprising program instructions for:

(a) determining sequence similarity between a database of polynucleotide sequences that correspond to accessible regions of cellular chromatin in a sample and one or more known sequences; and

(b) displaying information concerning the sequence similarity as determined in step (a).

121. A computer program product comprising a computer-useable medium and computer-readable program code encoded within the computer-useable medium, wherein the computer-readable program code

(a) comprises a database having a plurality of sequence records, wherein one or more of the sequence records include an identifier assigning that sequence record to one or more projects, wherein each project is based on determining whether the sequence record includes common and unique polynucleotide sequences corresponding to accessible regions of cellular chromatin; and

(b) effects the following steps with a computer system

(i) providing an interface that permits a user to query one or more projects;

(ii) locating sequence data corresponding to the query; and

(iii) displaying the sequence data corresponding to the query.

122. A computer program product comprising a computer-useable medium and computer-readable program code encoded within the computer-useable medium, wherein the computer-readable program code:

(a) comprises a database comprising a plurality of collections of polynucleotide sequences corresponding to accessible regions of cellular chromatin, different collections comprising accessible regions for different samples of cellular chromatin; and

(b) effects the following steps with a computer system:

(i) identifying sequences that are unique between collections selected by a user;

(ii) identifying sequences that are common between collections selected by a user; and

(iii) displaying common and unique sequences.

123. A computer program product comprising a computer-useable medium and computer-readable program code encoded within the computer-useable medium, wherein the computer-readable program code

(a) comprises a database comprising a collection of polynucleotide sequences corresponding to accessible regions of cellular chromatin in a chromatin sample; and

(b) effects the following steps with a computer system

(i) determining sequence similarity between two or more polynucleotide sequences selected by a user as compared to one or more known sequences; and

(ii) displaying the sequence similarity between the selected polynucleotides and known sequences.

124. The computer program produce of claim 123, wherein the one or more known sequences are genomic sequence of the organism from which the chromatin sample is obtained and wherein the displaying comprises indicating the location of the polynucleotide sequences relative to the genomic sequence.